

ABSTRACT OF THE DISCLOSURE

A method and system for reducing lexical ambiguity in an input stream are described. In one embodiment, the input stream is broken into tokens. The tokens are used to create a connection graph comprising a number of paths. Each of the paths is assigned a cost. At least one best path is defined based upon a corresponding cost to generate an output graph. The generated output graph is provided to reduce lexical ambiguity.